Training Interventions for Reducing Flight Mishaps

Gregg A. Montijo and David Kaiser
Crew Training International
Memphis, TN
gmontijo@cti-crm.com
dkaiser@cti-crm.com

V. Alan Spiker
Anacapa Sciences
Santa Barbara, CA
vaspiker@anacapasciences.com

Robert Nullmeyer
Air Force Research Laboratory
Mesa, AZ
robert.nullmeyer@af.mil

ABSTRACT

Increasing numbers of preventable mishaps across all military services led Secretary Rumsfeld and all Service Chiefs to call for a reduction in such events by 75% from 2003 levels. Most were attributed to human error. The highly task-loaded training and combat missions flown by fighter pilots place particularly high demands on effective management of cockpit resources for safe and successful mission accomplishment. While every flight training program already includes some form of resource management training, there is surprisingly little evidence regarding the effectiveness of varying training approaches to reduce flight mishaps.

This paper describes a project to help the Air Force reduce preventable mishaps by determining the specific root causes of fighter and unmanned aerial system mishaps, developing behaviorally-based training objectives, identifying promising training media alternatives, and defining specific measures of effectiveness. Mishap reports revealed several repeating problems in the areas of situation awareness, task management, and decision making in all platforms studied. A Delphi Panel of fighter, attack, and Predator pilots reviewed and in some cases, amplified the specific underlying human factors that are most challenging to pilots in tactical environments. The panel also considered the feasibility and probable value of nine potential training interventions. The Predator community was chosen for implementation and assessment of four interventions – focused academic training, interactive case histories, game-based multi-task practice, and a laptop-based simulator for team training. A review of historical Predator student records revealed that many trainees have difficulty mastering attention management, task prioritization, selecting a good course of action, and crew coordination.

Spiral implementation will enable the contributions of each intervention to be assessed using a controlled experimental design at an operational training unit. Anticipated benefits include increased student situation awareness, more effective task management, and improved decision making in subsequent flights, all contributing to the ultimate goal, fewer mishaps.

ABOUT THE AUTHORS

Gregg A. Montijo is the Combat Air Forces Crew/Cockpit Resource Management Program Manager for Crew Training International, Inc. He is a retired USAF command pilot with over 4,000 hours in the A-10, F-16, A-320 aircraft and various gliders. He earned a B.S. from the U.S. Air Force Academy in 1981, specializing in Human Factors Engineering. He also earned an M.S. in Procurement and Acquisition Management from Webster University in 1995. He holds an Airline Transport Pilot's license with a type rating in the B737 and has over 10 years program management experience at the Pentagon and in the civilian sector. He has been with Crew Training International since 2001.

David Kaiser is the Director of Courseware Development for Crew Training International and is responsible for all courseware development for both Crew Resource Management and Aircrew Training. He earned his master's degree from Embry Riddle Aeronautical University with a dual specialization in Aviation Education and Aviation Safety. A Navy aviator with combat experience, he is currently active in the Navy Reserve. He has managed various experimental projects and has been personally awarded two U.S. patents.

Alan Spiker is a principal scientist at Anacapa Sciences where he has worked since 1982. He received his PhD in experimental psychology from the University of New Mexico in 1978. At Anacapa, he is responsible for managing human factors research projects in advanced technology, and has specialized in aviation training and human performance for all branches of service and the airline industry. During the past 10 years, he has conducted research

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 DEC 2008**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **Training Interventions for Reducing Flight Mishaps**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Crew Training International Memphis, TN**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release, distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**12**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

on the cognitive underpinnings of effective, safe aircraft operation, including crew resource management, multi-tasking, planning, and decision-making.

Robert Nullmeyer is a research psychologist with the Air Force Research Laboratory at Mesa, AZ, where he has been since 1981. He has conducted research on training system evaluation, simulator training effectiveness, and crew resource management training needs analysis. He earned a PhD in Experimental Psychology from the University of New Mexico in 1979.

Training Interventions for Reducing Flight Mishaps

Gregg A. Montijo and David Kaiser
Crew Training International
Memphis, TN
gmontijo@cti-crm.com
dkaiser@cti-crm.com

V. Alan Spiker
Anacapa Sciences
Santa Barbara, CA
vaspiker@anacapasciences.com

Robert Nullmeyer
Air Force Research Laboratory
Mesa, AZ
robert.nullmeyer@af.mil

INTRODUCTION

The central role of human error in flight mishaps is well documented. Helmreich and Fouchee (1993) reported that flight crew actions were causal in more than 70% of worldwide air carrier accidents from 1959 to 1989 involving aircraft damaged beyond repair. In commercial aviation, mishaps attributed to human error appear to be declining. Shappell, Detwiler, Holcomb, Hackworth, Boquet, and Wiegmann (2006) reported a steady decline in percentages of commercial aviation accidents in which human error was causal from 73% in the early 1990s to less than 60% in 2000-2002. Similarly, Baker, Qaing, Rebok, and Li (2008) reported a drop in air carrier mishaps involving human error from 42% in the 1980s to 25% in 1998-2002.

In contrast, mishap rates rose slightly but steadily from 1999 through 2003 in all U.S. military services following decades of improvement. In the Air Force, Luna (2001) reported that human factors were causal or major contributors in over 60% of Class A mishaps from FY1991 through FY2000. Heupel, Hughes, Musselman, and Dopslaf (2007) reported similar percentages in Air Force mishaps from FY2000-FY2006 (64%). Rising mishap rates across all military services led to directives from Secretary Rumsfeld to reduce preventable mishaps (Rumsfeld, 2003, 2006). This, in turn, generated pledges from all Service Chiefs of Staff to reduce preventable mishaps by 75% from 2003 levels. The U.S. Coast Guard (2008) compared 2007 Class A flight mishap rates across all military services. Relative to mishap rates in the preceding four years, some organizations showed more progress toward reducing mishaps than did others. The Navy and Marine Corps reduced mishap rates by about one third in 2007 relative to the previous four years. The Coast guard had no Class A mishaps in 2007. In contrast, mishap rates in 2007 increased slightly service-wide in the Air Force and Army compared to the previous four years.

Further analyses of FY2007 Air Force Class A mishaps revealed unusually high numbers of F-15 and F-16 mishaps. There were six F-15 Class A mishaps in 2007 versus 2.8 per year from FY2003-FY2006. Two were attributed to human error. Thirteen F-16 mishaps rose from 6.8 historically. Seven involved human factors. Predator mishap counts rose slightly to five in 2007 from an historic average of 4.5. Three involved human factors. These three platforms accounted for 80% of all Air Force Class A mishaps in 2007, and half of these were attributed to human factors.

In light of enviable reductions in human factors-related commercial aviation mishaps, it may be useful to review safety training practices in that arena. Helmreich, Merritt, and Wilhelm (1999) documented a progression of crew resource management (CRM) training philosophies and goals through four distinct generations. They concluded that the original safety-related goals of CRM appeared to have become lost over time and proposed a fifth generation of CRM training explicitly focused on error management. Five data sources were recommended to sharpen that focus: (a) formal evaluations of flight crews, (b) incident reports from aviators, (c) surveys of flight crew perceptions regarding safety and human factors, (d) parameters of flight from flight data recorders, and (e) line operations safety audits (LOSA). Each illuminates a different aspect of flight operations.

Helmreich, Wilhelm, Klinect, and Merritt, (2001) studied threats to safety and the nature of errors in three airlines using LOSAs. Striking differences were observed among these air carriers regarding both threats to safety and the nature of operator errors. Based on this experience, Helmreich and his colleagues concluded that individual air carriers cannot assume their training requirements will correspond to normative data from the industry. Rather, they postulated that organizations must have current and accurate data regarding the true nature of threats and errors to shape effective training content and structure assessments of training impacts. They proposed a sixth generation of CRM training that adds the need to understand an organization's threats to safety to the previous domain of error management.

We believe that threats to safety in military operations need to be better understood and error reduction training needs to be more focused if the military is to achieve the desired reductions in preventable mishaps that have been enjoyed by their commercial counterparts. To that end, several analyses of Air Force mishap data were recently completed. Nullmeyer, Stella, Montijo and Harden (2005) analyzed attack, fighter, and tactical airlift mishaps, and Nullmeyer, Herz, Montijo and Leonik (2007) investigated Predator mishaps. Both reconnaissance

(RQ-1) and multi-mission (MQ-1) platforms were included in the Predator analyses. Three skill areas were consistently cited as factors in Air Force fighter, attack, and Predator flight mishaps: (a) situational awareness development and maintenance, (b) task management, and (c) decision making.

We recognize that mishap reports are not sufficient by themselves to structure training. Dekker (2003) described several potential problems associated with over-reliance on human error taxonomies, including risks associated with removing the context that helped produce the error. Such concerns imply that quantitative mishap human factors trends must be viewed in the context of other information to develop truly robust training interventions that are likely to impact safety and effectiveness. To that end, we augmented the safety data with expert opinion and trends in student records.

The remainder of this paper describes a project that intends to help the Air Force reduce preventable mishaps by determining the particular human factors skills that are most relevant to the fighter and unmanned aerial vehicle (UAV) communities, identifying several potential strategies for reducing subsequent operator error through training, and developing a concept of operations to test the effectiveness of the most promising training interventions that would address deficient skills.

AIR FORCE CLASS A MISHAPS

The first step in this project was to identify current human factors deficiencies in high-workload fighter and UAV tactical environments. To accomplish this, we reviewed reports of A-10, F-15C, F-15E, F-16, and RQ-1/MQ-1 Class A mishaps ($1 million damage or fatality) from FY1996 through FY2006. The Air Force Safety Center (AFSC) documents Class A mishaps in a variety of forms, and our analyses combined information from several of these sources. The first was a detailed **Human Factors Database** populated and maintained by AFSC Life Sciences Division analysts. This database lists all factors cited regarding the roles played by operators, maintainers, and other personnel in each mishap. Our research team created aircraft-specific databases to facilitate identification of trends and idiosyncratic results. We further investigated specific causes and contributing factors using more detailed mishap source documents, primarily the **Safety Investigation Board Report** and the **Life Sciences Report**. Qualitative analyses of discussions in these reports were accomplished to gain a better understanding of the underlying behaviors that led to each element being cited.

The AFSC Human Factors Database listed all human factors cited in the Life Sciences Report section of each full mishap report and provides a contribution score: 4=causal, 3=major factor, 2=minor factor, 1=minimal factor and 0=present, but not a factor for mishaps through FY2006. From this database, we created a combined index (frequency and importance) by summing these scores across mishaps for each cited human factors element. These weighted sums were then used to rank-order the individual elements, with a separate ranking created for each weapon system. The top ten causal and major contributing factors cited in the Human Factors Database across the platforms addressed in this study are shown in Table 1. Numbers of mishaps by weapon system are listed immediately beneath each aircraft type. For example, there were 20 A-10 mishaps. The remainder of the table shows the numbers of mishaps in which a specific human factors element was cited as a causal or major factor. In the 20 A-10 mishaps, channelized attention was cited in nine mishaps, and task misprioritization was cited in seven.

Channelized attention, task misprioritization, and *selecting the wrong course of action* were cited as problems in every platform analyzed. Factors beyond these top ten were also cited, but usually in only one

Table 1. Top Ten Root Causes in Tactical Aircraft Class A Mishaps (FY1996-FY2006)

Aircraft Type (Numbers of mishaps) **Human Factors Elements:**	**A-10** (20)	**F-15C** (14)	**F-15E** (9)	**F-16** (86)	**RQ-1/MQ-1** (30)
Channelized attention	9	8	3	25	8
Task misprioritization	7	3	2	17	4
Misperception	4	4		14	
Selecting wrong course of action	3	3	1	9	4
Wrong technique/procedure		6	1	8	4
Cognitive task oversaturation	5	3		10	
Spatial disorientation	3		2	11	
Risk assessment				11	3
Distraction/inattention			3	7	3
Inadequate in-flight analysis	7	2	1		2

or two platform types. Necessary action delayed and event proficiency were problematic in A-10 mishaps. Crew coordination, checklist error, confusion, inadequate written procedures, and interface design issues were commonly cited in Predator mishap reports. These quantitative analyses suggest that a number of threats to safety are common across fighter, attack, and reconnaissance platforms, but there are a number of platform unique issues as well, particularly for Predator operators.

CANDIDATE TRAINING INTERVENTIONS

Through reviews of Web planning-related sites, technical descriptions of interventions in the literature, and discussions with training analysts, nine promising candidate training interventions were identified that would address the skills emerging from the mishap analyses. The interventions spanned the spectrum of possible solutions from self-study and focused academics to specialized simulation and network technologies. We defined a "promising" intervention as one that has a potentially positive impact on one or more of the HF skill deficiencies identified, is logistically and technologically compatible with a mission-oriented training environment, and is feasible for implementation in this Phase II Small Business Innovative Research project, (i.e., can be implemented and evaluated within program time and budget constraints [2 years and $750,000]). The interventions were not necessarily mutually exclusive, and could, as needed, be bundled into a more comprehensive intervention "package." The nine identified candidate training interventions are shown in Table 2.

Table 2: Potential Training Interventions

Intervention	Description	Example
Self Study	Material is presented to the aircrews in text format via e-Learning to study at their own pace.	Chair Fly or Table Top a Mission - Warfighter might review choke points in a mission during pre-flight and think through courses of action that could be taken to reduce workload ahead of time.
Classroom-Style Training	Material is presented via a number of delivery styles: • Pure lecture • Guided lecture and discussions • Facilitated lecture (guided learning) • Facilitated lecture with in-class exercises • Computer-based self-study, plus facilitated advanced in-class interactive case studies/exercises	Videos could be taken of successful and unsuccessful crews performing the HF skill of interest in the mission trainer. To enhance instruction, the videos could be "scripted," using role-playing instructors, to highlight particular HF positive or negative behaviors.
Computer-Based Training	Training can be provided in specific skills, where a background scenario could be given to "draw" the warfighter into the context.	The team trainer GemaSim - Crews are given academics to understand their individual reactions to stress, how to recognize stress limits of others, and how to function effectively as a team under stress. Crews are assigned to a laptop-based network to complete a mission (space) exploration in which they compete against other teams of crewmembers. During the mission they are subjected to stress in order to experience breakdown in cognitive capabilities. Crews are observed and debriefed on their experience.
Part Task Trainer	A moderate fidelity simulator could be designed that has high fidelity for the HF skill of interest, with lower fidelity for other parts of the mission. • Specially designed equipment • Existing equipment with specific software or mission profiles	A CRM Part Task Trainer (PTT) was developed for the C-130 community that had fully functioning radios so copilots and navigators could learn to communicate during airdrops. The rest of the simulator – flight controls, visuals, multi-function display – was of lower fidelity, just enough to support the aircrew for the other parts of the mission.

Gaming Solution	CBT instructional material transformed into a game where points are awarded, repetitive play is encouraged, and competition is emphasized by displaying the top scores.	Game requiring players to monitor and respond to several simulations of UAV displays (e.g., heads-up display screens, chat lines, imagery, map, etc.)
Full Mission Trainers	Correct skill deficiencies in a Full Mission Trainer environment • Add software to existing system • Modify mission profile to train skill	Simulators can be configured that have fairly high fidelity to support multi-crew teamwork training in customized scenarios. Problem HF skills can be addressed through repetitive practice, feedback, and debriefs.
Dedicated Mission Trainers	Simulator training specifically dedicated to specific skills tied to safety of flight • Use existing simulators and modify software to train specific CRM skills • Relies heavily on debriefing	Simulators that emphasize particular missions can be used where the targeted HF skills are a major player for that mission. (e.g., channelized Attention could be selected for highlighting training in the context of air/ground missions with visually complex enemy laydowns).
Modify Existing Simulator Profiles	Use existing training capabilities, insert specific training events that would stress and target particular HF skills. • Requires in-depth analysis of existing profiles • Specific mission events are needed to have desired behavioral outcomes • "The Gouge" can quickly develop among flight crew - negates training • Easiest in terms of schedule, cost	A particular training profile could be modified by inserting additional task stressors, (e.g., threat pop-ups, reduced visibility, caution lights, etc.), to provide training in task prioritization. Embedded performance standards would be included in the events, as well as feedback provided in the debrief.
Networked Solutions	Full spectrum missions flown in simulators linked with other participants • May be stand-alone in nature or part of a Joint exercise. • May blend real world and synthetic environments. • The ability to capture individual behavior in a dynamic computer environment with a wide-range of possible outcomes is a potential challenge	Distributed Mission Training (DMT)/ Distributed Mission Operations (DMO)

THE DELPHI PANEL

A Delphi Panel of F-15, F-16, A-10, and RQ-1/MQ-1 warfighter experts was convened to solicit their opinions on skill deficiencies and potential training interventions. To accomplish these goals, we constructed a multi-faceted instrument designed to collect both quantitative data regarding problem frequency and difficulty, and qualitative data reflecting the panel's comments regarding key problems, issues, and explanations. As such, the instrument was consistent with the project's multi-method, multi-measure approach to identifying, defining, measuring, and evaluating high-payoff CRM skills. Because of high Operations Tempo (OPTEMPO) and scheduling issues, we restricted our panel to a half-day at the U.S. Air Force Weapons School, Nellis AFB, NV. This location permitted at least one representative from each of the aforementioned weapon systems to attend, with the Predator community supplying three people. Thus, a total of six experts attended the three-hour session. Despite the logistical problems in convening the panel, the qualifications and experience levels of the participants were impressive. All were officers, O-4 and above, with most having hundreds or thousands of hours operational training and combat experience with their particular weapon system. All participants were highly-motivated to support the present project, and each appeared to be genuinely interested in improving CRM skills for their weapon system. In short, the panel composition and tone was ideal for our purposes.

Identifying Human Factors Skills

Panel members were given a list of skills that had been derived from the Class A mishap reports. The list included 19 skills – the ten factors listed in Table 1 plus nine others. In each case, panel members were asked to rate each skill using the following five-point scale:

1. No problems in training/operational missions
2. Minor problems in training/operational missions
3. Some problems in training/operational missions
4. Major problems in training/operational missions
5. Severe problems in training/operational missions

Panel participants reviewed each skill in turn, providing a rating and, in some cases, offering written comments explaining the basis for their ratings. A moderated discussion concerning issues and problems regarding these skills followed.

The initial series of analyses was performed on the data from the six panelists who represented all four tactical weapons systems (three of the six panelists were Predator operators). Table 3 summarizes the mean importance/problem ratings for the skills that were identified in the mishap report analyses. They are presented in descending order of mean rating, where the scale can range from 5 (severe problem) to 1 (no problem). The top four skills based on mishap reports are indicated in red italics. To provide a metric for making comparisons, we computed the variance of ratings within each skill, took the average, and then computed the average standard error about the mean. Doubling that number provides a good estimate of the typical rating difference that would be considered statistically significant if inferential tests were conducted (Hays, 1973). Our analysis showed this value to be about .75. For example, on the basis of this metric, we could conclude that the average rating for Cognitive Task Oversaturation (3.7) is statistically higher than Task Misprioritization (2.9). While not used to completely guide our analyses or interpretations, such an index should be kept in mind when attempting to draw firm conclusions from an admittedly small sample size.

The quality of the information provided, given the high experience levels of the panelists, more than compensates for the lack of statistical power in any test that one would conduct. It is evident from the table that although the top four human factors topics from mishap trends are, by and large, among the higher-rated problems, there are others that the experts elevated in terms of relative importance. In particular, Cognitive Task Oversaturation was the factor that was rated as being most problematic by the Delphi Panel, even though it did not occupy that spot in any platform based on mishap report analyses, and was not cited at all in Predator mishap reports. This element refers to the magnitude or variety of inputs exceeding operator limitations to process information.

Table 3. Mean Rating of Importance/Problem for 19 Human Factors

Human Factor	**Mean Rating** (5=max, 1=min)
Cognitive Task Oversaturation	3.7
Channelized Attention	*3.4*
Inadvertent Operation	3.3
Inadequate In-flight Analysis	3.0
Confusion	3.0
Wrong Course of Action Selected	*3.0*
Task Misprioritization	*2.9*
Crew Coordination Breakdown	2.9
Misperception of Speed, Distance, Altitude	*2.8*
Wrong Technique	2.6
Distraction	2.5
Limited Systems Knowledge	2.4
Poor Intracockpit Communication	2.4
Checklist Error	2.3
Inattention	2.2
Complacency	2.2
Subordinate Style	2.0
Overcommitment	2.0
Poor Risk Assessment	1.8

Note: All four tactical weapon systems are included.

Inadvertent Operation reflects a poor choice of switch or function operation, which is especially problematic with the software intensive Predator operator console. Inadequate Inflight Analysis and Confusion are problem areas that appear as factors in multiple systems.

Selecting Training Interventions

The Delphi session then turned to candidate training interventions. The research team explained the nine different training interventions the panel would be asked to consider, corresponding to the ones listed in Table 2. The interventions were presented in reverse order of fidelity, beginning with self-study, followed by classroom-style training, computer-based solutions, full mission trainers, dedicated mission trainers, modification of existing simulator profiles, and networked solutions. Note that these interventions are actually *categories* of technologies that span a spectrum of possible solutions to the HF skills problems provided in the first part of the Delphi session. The presentation was interactive, with panel members asking questions and offering suggestions. Two ratings were asked of each of the nine interventions. The first was a five-point, behaviorally-anchored scale that had participants rate the intervention's estimated degree of impact on the targeted human factors skills. A second five-point scale called for rating the feasibility of implementing the intervention in an operational training squadron. Besides the rating, the instruments contained space for panel members to make amplifying comments; free-flowing discussions followed the rating process.

During the Delphi Panel session, one of the panel members, the commander of the 11th Reconnaissance Squadron (RS), indicated his desire to have other members of his squadron review the instrument and provide their assessment. The commander's endorsement of the project, and his willingness to have the MQ-1 Predator community serve as claimants, was unquestionably a turning point in the project. Per the commander's suggestions, we supplied the squadron with additional copies of the instruments. Several weeks after the workshop, three additional completed instruments were provided to the project team. It was at this point that we decided to perform two analyses. The first was on data from the six original Delphi Panel members. The second was on the six MQ-1 operators, three from the Delphi session and three survey respondents from the 11th RS, who comprised our sample. The SMEs from the other platforms provided highly similar ratings, so only the ratings from the six MQ-1 operators are shown in Table 4. The left part of the table summarizes the mean ratings of expected impact in descending order; the right portion provides the average ratings for intervention feasibility.

As can be seen, there is a marked divergence between the two sets of ratings. The interventions that panel participants rated as having the highest impact were mostly associated with being the least feasible to implement, and vice versa. Analysis of the comment data provides some ready explanations for these results. In this regard, full mission trainers were clearly seen as an effective way to train many human factors skills. Unfortunately, their feasibility for implementation within the time and resource constraints of this project is limited. Conversely, computer-based training, which was summarily dismissed by attendees based on recent negative experience, was rated poorest on impact yet was recognized for being quite feasible. It should be noted that classroom training, the clear favorite for feasibility, also received respectable marks for potential impact. This bodes well for attempts to improve error reduction via classroom training by targeting specific human factors skills with new case examples and highly focused spin-up training. This issue is taken up later in the paper when we discuss the interventions chosen for implementation.

Table 4. Mean Ratings of Intervention Impact and Feasibility (RQ-1/MQ-1 only)

Intervention Impact		Intervention Feasibility	
Intervention	**Mean Rating**	**Intervention**	**Mean Rating**
Full Mission Trainer	4.3	Classroom Training	3.8
Classroom	4.2	Computer Based Training	3.3
Dedicated Mission Trainer	4.1	Handheld Game	3.3
Modify Existing Simulator	3.8	Self Study	3.2
Self Study	3.6	Network Solutions	3.0
Part Task Trainer	3.5	Part Task Trainer	2.5
Network Solutions	3.2	Full Mission Trainer	2.5
Handheld Game	3.0	Dedicated Mission Trainer	2.4
Computer-Based Training	2.7	Modify Existing Simulator	2.2

Finally, we received the endorsement of the 11th RS Commander to host field studies of resulting training interventions. Having an operational claimant who eagerly awaits our interventions ("I would like them today!") is a reaction that is all-too-rare in the research and development community. As we describe below, we plan to work extremely closely with the 11th RS Commander and his organization to ensure that the training interventions we specify, prototype, develop, and implement meet the squadron's current and projected training requirements.

TRAINING RECORDS ANALYSIS

With the selection of the Predator training program as the environment in which interventions would be implemented and evaluated, training records in this community were analyzed to identify tasks that are particularly difficult or challenging for students, conducting both quantitative analyses on grades and content analyses on instructor comments.

Records from 70 student pilots and 75 sensor operators were reviewed from the Predator Operator Basic and Requalification course, focusing on student performance in the final 2 flying training sessions preceding the checkride. Instructors used a 5-point grading scale from 0 to 4, with a "2" or higher representing a passing level of performance. No "0" scores were observed, but 101 "1s" were recorded for pilots and 62 "1s" for sensor operators. These less-than-passing grades at the end of training were concentrated in 7 of the 45 graded pilot task elements and 4 of the 50 sensor operator task elements.

For pilots, the task elements were:

- Buddy laze procedures
- Launch
- Target acquisition, aircraft position
- Operational mission procedures
- Deconfliction plan/execution
- AGM-114 employment
- Airmanship/aircraft control

For Sensor operators, the task elements were:

- Launch
- Mission CRM/crew coordination
- Mission planning/preparation
- AGM-114 employment

These problematic task elements were further analyzed with the aid of instructors to identify common underlying skill areas. Four skill areas emerged: avoiding channelized attention, Prioritizing tasks, selecting an appropriate course of action, and crew coordination. Two particularly challenging syllabus events were also identified that require students to apply these skills: a simulator-based emergency procedures scenario, and a flightline tactical mission that occurs shortly before the final checkride.

TRAINING INTERVENTIONS SELECTED

To accelerate skill development in the problem areas that emerged from the preceding activities, four training interventions were selected for further development and evaluation: enhanced focus academic training; interactive, web-based or desktop case histories; gaming computer-based training to develop individual task monitoring and task management skills; and a computer-based team training environment. Each is further described below.

Enhanced focus academic training is based on the foundations of adult learning principles. These principles are presented in a facilitation style, in contrast to lecture style, in order to actively engage the following androgological principles (Knowles 1980; Knowles, Holton & Swanson 1998): (a) fulfilling the learner's need to know (helping students see the value of training and how it applies to them in their job); (b) allowing students to be more self-directed; (c) leveraging a variety of experiences to build on some learners' already-acquired experiences, transferring that knowledge base to those who have less experience; and (d) specifically designing the learners' experience to increase their readiness, orientation, and motivation to learn.

Interactive, web-based or desktop case history is based on a computer-based training system developed for the Navy that took articles from the Navy's *Approach* magazine, added supplemental information to reinforce core concepts in human performance disciplines, and presented this information in electronic form (Spiker, Hunt, and Walls, 2005). It was intended for use as an adjunct to classroom instruction. The summaries are written in a readable style designed to both entertain and educate. The case study is followed by a short set of fairly difficult questions that are written to require the student to read the case study and understand the main points. It was clear from the Delphi Panel that our experts all had less-than-stellar experiences with CBT in the past. The prevailing view was that much of what they had experienced was merely "electronic page turning," and not particularly engaging. In recognition of this, the intent with this medium is to develop compelling, interesting, informative, and memorable instruction by design.

Computer-based gaming of individual skills as an intervention is loosely adapted from a test of multi-tasking ability called *SYNWIN* (Elsmore, 1994). While SYNWIN's prior use has been as a selection test, our plan calls for casting the concept in a game format that can be played by trainees while they are receiving their initial CRM training. Our belief is that promoting the instructional material in the form of game, where scores can be competitively acquired and even posted, will overcome some of the negative reaction to CBT that was discussed in the previous task. The test requires users to simultaneously monitor four quadrants of the primary display screen. The upper left quadrant of the screen displays a letter recall task in which participants click a button to indicate whether a probe letter was a member of a previously displayed set of letters (the subject must remember that set of letters). The upper right quadrant presents an arithmetic task, where participants solve simple, randomly-generated three-digit addition problems. A visual monitoring task is in the lower left, where participants click on a gauge to reset a slowly moving pointer before it reaches the zero mark. The lower right quadrant has an auditory monitoring task where participants listen to a series of high and low frequency tones, and click a button when they hear a high frequency tone.

From an instructional perspective, one of the strongest features of games is that they offer ample opportunity for practice and repetition. As well, games usually provide immediate, clear feedback and require criterion skill mastery to move to the next level. But the most-cited advantage of using game elements in instruction is the motivational factor – people usually want to play games and will voluntarily devote a great deal of time to mastering the skills and rules of the game. This may be particularly relevant with many of today's students and trainees who, as digital natives, have been raised in a technology-dominated environment, with hours of video and computer game playing.

Besides transforming the SYNWIN test concept into a game, we will also explore altering each of the four tasks so they have more in common with tasks that UAV operators presently perform. For example, the memory recall task, which in SYNWIN consists of random letter/number strings, can be converted into a more meaningful task where the aviator is to recall sequences of letters and numbers that might correspond to airfield designations, waypoints, landmarks, navigation aids, etc. While the cognitive task – holding information in memory for an extended time – is the same, the actual task will more resemble what is actually required of Predator pilots and sensor operators. Similarly, the addition task could be expanded to include other mental operations that UAV operators must perform, such as doing basic geometry to compute descent angles, calculating distance between waypoints, or extrapolating airspeeds and leg times, to name a few. Similarly, the visual monitoring task does not have to be restricted to a fuel gauge. It too can be altered to more closely mimic UAV operations. For example, we could use an embedded video (say, from a sensor) and ask the subject to monitor it for some dynamic characteristic (e.g., a target).

Computer-based team training is designed to exercise team functions and behavior in a stressful environment. The GemaSim team trainer (Figure 1) allows for the experience, observation, analysis, modification and consolidation of authentic behavioral patterns that emerge under stressful conditions. Once under stress, humans may switch from established norms, industry practice, etc. and apply a different set of dominant logic pathways, resulting in abnormal behaviors. This effect has been observed in such high-risk/high-pressure industries as aviation, rail, medicine and executive management. The intent of this device is analogous to the high altitude chamber training where pilots, although taught the effects of hypoxia, all experience different symptoms. Similarly GemaSim provides an enjoyable, but serious and relevant simulation activity that allows for one's own

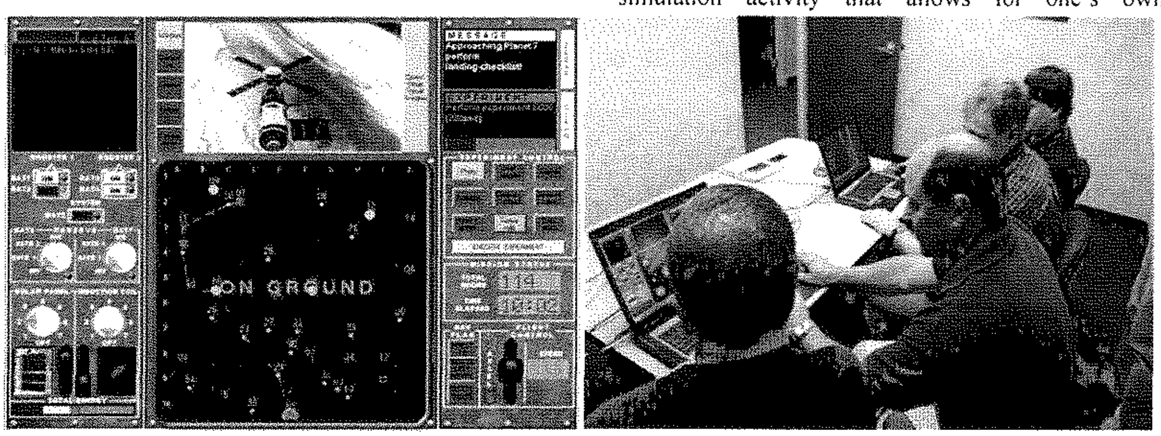

Figure 1: **Students under stress during GemaSim team training**

behavioral patterns to be experienced, together with those of a specific team under situations of increased pressure. Through an understanding of the causal factors of human behavior, and by analysis of one's own behavioral patterns, these can be modified, re-exercised and consolidated.

IMPACT ASSESSMENT

Our plans call for conducting an 18-month assessment of the four training interventions at Creech AFB. We plan to follow Kirkpatrick's (1996) four-level evaluation approach in which data are collected to assess: (a) the reaction of trainees to the **usability and usefulness** of the training intervention (Level I); (b) the **amount of learning or skill acquisition that occurs** from the training (Level II); (c) **if the skills that are trained transfer to the job** (flight) environment (Level III); and (d) the **benefits that accrue to the organization** as a result of the training (Level IV).

As Salas and his colleagues have noted (Salas, Fowlkes, Stout, Milanovich, & Prince, 1999), few studies of the overall effectiveness of CRM training (Level III) have been conducted, and even fewer assessed all four levels in the same context. We plan to fill this empirical data gap by implementing a series of measures at various points in the training curriculum, including a baseline period before the four interventions are introduced. A new class of pilot and sensor operator training is offered roughly every 3 weeks at the squadron, with some 20 students attending per class. Importantly, we will be performing a fairly controlled evaluation as only half the classes will receive the training interventions, with the other half serving as a control (receiving only traditional CRM). The large sample size should give us sufficient statistical power to perform multivariate analysis of variance and follow-up test procedures.

Our training interventions will be incorporated into the current curriculum as a series of four "spirals" in order to restrict our footprint on on-going operations and to help manage the complexities of parallel development. The first spiral will consist of only the first intervention (focused academics). The second spiral will entail implementing focused academics and interactive case histories. Spiral 3 will consist of the first two interventions plus the game-based training. The final spiral will comprise all four interventions. Each spiral will be implemented in two classes (about 40 students per condition), where another two classes will serve as a control. This design will let us gauge both the training effectiveness of the overall intervention package (relative to current CRM training), as well as the contributions of the individual interventions to effectiveness.

To measure intervention impact, we will employ a cadre of specialized instruments and review the squadron's regular training records. First, we will insert questions into the end-of-course critique to assess student reaction to the training in the four HF skills of interest (Level I assessment). Second, we will conclude each intervention with a comprehension assessment to ensure that learning of the HF skills has occurred (Level II).

Instructors and observers will use a specialized gradesheet to measure proficiency in the simulator training sessions following the interventions. These sessions will give us the much-needed Level III data to gauge whether the skills we believe students have learned in our training interventions actually manifest themselves in realistic flight conditions. This gradesheet will consist of some half-dozen key behaviors associated with each HF skill. For example, the HF skill "avoids channelized attention" would be decomposed into such key behaviors as: effective cross-check includes all relevant displays; cross-check does not stagnate; switches attention as the situation priority changes; etc. Importantly, key behaviors will be defined to support reliable observation by instructors and raters.

CONCLUSION

Our main purpose in this project is to help reduce preventable flight mishaps, so our assessment of benefits to the organization needs to address the impact of these interventions on safety of flight. A direct assessment of that effect will require longitudinal tracking of Predator crews beyond the time frame of this project. This project will, however, determine the ability of our interventions to accelerate the development of skills that were lacking in previous Class A mishaps.

Much of what we learned to date is encouraging. The vast majority of Air Force Class A mishaps (78%) in 2007 involved F-15, F-16, and Predator operations, and the root causes of mishaps in these three platforms have much in common—mishap reports from all three communities frequently cite channelized attention, task misprioritization, and course of action selected. Our panel of experts from each of these systems added cognitive task oversaturation as a fourth problem area. As a result, it appears that a finite set of factors is driving Air Force preventable Class A mishaps.

Our approach assumes that these problem areas reflect trainable skills. Given the support that we enjoy with the Predator community, this project represents an excellent opportunity to move from problem statements to validated solutions. Interventions that positively impact on subsequent attention and task management or improved decision making for Predator crews should be directly applicable to the fighter and attack communities.

REFERENCES

Baker, S.P., Qiang, Y., Rebok, G.W., and Li, G. (2008). Pilot error in air carrier mishaps: longitudinal trends among 558 reports, 1982-2002. *Aviation, Space, and Environmental Medicine*, 79(1), 2-6.

Dekker, S.W.A. (2003). Illusions of explanation: A critical essay on error classification. *The International Journal of Aviation Psychology, 13*(2), 95-106.

Elsmore, T. (1994). SYNWORK1: A PC-based tool for assessment of performance in a simulated work environment. *Behavior Research Methods, Instruments & Computers, 26,* 421-426.

Hays, W.L. (1973). *Statistics for the social sciences* (2nd ed.). New York: Holt, Rinehart, & Winston.

Helmreich, R.L. & Foushee, H.C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E.L. Wiener, B.G. Kanki, and R.L. Helmreich (Eds.), *Cockpit resource management*, San Diego, CA: Academic Press.

Helmreich, R.L., Merritt, AC., & Wilhelm, J.A. (1999). The evolution of crew resource management training. *The International Journal of Aviation Psychology, 9*(1), 19-32.

Helmreich, R.L., Wilhelm, J.A., Klinect, J.R., & Merritt, A.C. (2001). Culture, error, and crew resource management. In E. Salas, C.A. Bowers, & E.Edens (Eds.), *Applying resource management in organizations: A guide for professionals. Hillsdale,* NJ: Erlbaum.

Heupel, K. A., Hughes, T.G., Musselman, B.T., & Dopslaf, E.R. (2007). *USAF aviation safety: FY 2006 in review.* http://safety.kirtland.af.mil

Kirkpatrick, D.L. (1996, January). Techniques for evaluating training programs. *Training & Development.*

Knowles, M.S. (1980). *The Modern Practice of Adult Education: From Andragogy to Pedagogy.* Englewood Cliffs, NJ: Cambridge Adult Education.

Knowles, M.S., Holton, E.F., III, &. Swanson, R.A. (1998). *The Adult Learner.* Houston: Gulf Publishing.

Luna, T.D. (2001). *USAF aviation safety: FY 2000 in review.* http://safety.kirtland.af.mil

Nullmeyer, R.T., Stella, D., Montijo, G.A., & Harden, S.W. (2005). Human factors in Air Force flight mishaps: Implications for change. *Proceedings of the 26th Interservice/Industry Training Systems and Education Conference.* Orlando, FL.

Nullmeyer, R.T., Herz, R.A., Montijo, G.A., & Leonik, R. (2007) Birds of Prey: Training solutions to human factors problems. *Proceedings of the 28th Interservice/Industry Training Systems and Education Conference.* Orlando, FL.

Rumsfeld, D. (2003). *Reducing Preventable Accidents.* Memorandum for Secretaries of the Military Departments. The Secretary of Defense, Washington DC.

Rumsfeld, D. (2006). *Reducing Preventable Accidents.* Memorandum for Secretaries of the Military Departments. The Secretary of Defense, Washington DC.

Salas, E., Fowlkes, J.E., Stout, R.J., Milanovich, D.M., & Prince, C. (1999). Does CRM training improve teamwork skills in the cockpit? Two evaluation studies. *Human Factors, 41*(2), 326-343.

Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. (2006). Human error and commercial aviation accidents: A comprehensive, fine-grained analysis using HFACS (DOT/FAA/AM-06/18). Washington, D.C.

Spiker, V.A., Hunt, S.K., & Walls, W.F. (2005). *User reaction to annotated approach* (CNAP Technical Memorandum). San Diego, CA: Commander Naval Air Force, US Pacific Fleet.

U.S. Coast Guard (2008). *FY07aviation safety report,* available at http://www.uscg.mil/safety.